

CLAIMS

What is claimed is:

1. A method of analyzing a multi-dimensional data set to
5 determine the presence of one or more peaks within the data set,
the peaks being representative of respective compounds, used for
detecting at least one constituent compound in a sample mixture,
comprising the steps of:

generating a first multi-dimensional data set, the first
10 data set being representable by a first data array including a
first dimension corresponding to compound separation and a second
dimension corresponding to compound characteristics;

analyzing the first data array in the compound separation
dimension to generate one or more original representations of the
15 compound separation, the original representations including one or
more peaks and one or more regions having no peaks;

determining noise characteristics of the first data set
based on an analysis of the regions of the original
representations having no peaks;

20 reducing noise in the multi-dimensional data set by
performing matched filtration of the original representations of
the compound separation with at least one transfer function based
on the determined noise characteristics, thereby generating one or
more noise-reduced representations of the compound separation, the
25 noise-reduced representations including one or more peaks;

generating a second multi-dimensional data set with reduced
noise based on the noise-reduced representations, the second data
set being representable by a second data array including a
compound separation dimension and a compound characteristics
30 dimension; and

analyzing the second data array in the compound separation
dimension and the compound characteristics dimension to determine
the presence of one or more peaks within the second data set.

2. The method of claim 1 wherein the original representations of the compound separation generated in the first analyzing step and the noise-reduced representations of the compound separation generated in the noise reduction step are selected from the group consisting of an extracted ion chromatogram, an electrochromatogram, and an electropherogram.

3. The method of claim 1 wherein the peaks representative of respective compounds are selected from the group consisting of chromatographic peaks, electro-chromatographic peaks, and electrophoretic peaks.

4. The method of claim 1 wherein the first data set generated in the first generating step and the second data set generated in the second generating step comprise respective liquid chromatography-mass spectrometry (LC-MS) data sets, the first and second LC-MS data sets being representable by respective first and second data arrays, each respective data array including a chromatographic time dimension and a mass spectral dimension, and wherein the original and noise-reduced representations of the compound separation comprise respective original and noise-reduced extracted ion chromatograms.

5. The method of claim 4 wherein the determining step includes determining the noise characteristics of the first data set, the noise characteristics being representable by a power density spectrum of the regions of the original extracted ion chromatograms having no chromatographic peaks.

6. The method of claim 5 wherein the determining step includes determining the noise characteristics of the first data set, the noise characteristics being representable by a plurality of power

density spectra of respective regions of the original extracted ion chromatograms having no chromatographic peaks, and averaging the plurality of power density spectra.

- 5 7. The method of claim 5 wherein the noise reduction step includes determining at least one transfer function for matched filtration according to the equation

$$H(f) = S^*(f)/P_{NN}(f),$$

10 $S^*(f)$ being the complex conjugate of the Fourier transform of the function representing an expected shape of the chromatographic peaks, and $P_{NN}(f)$ being a power density spectrum.

- 15 8. The method of claim 7 wherein the expected shape of the chromatographic peaks is Gaussian.

9. The method of claim 7 wherein an expected width of the chromatographic peaks is within a predetermined range of peak
20 widths.

10. The method of claim 7 wherein the determining step includes determining at least one transfer function for matched filtration, wherein
25

$$P_{NN}(f) = \int R_{NN}(t) \exp(-j2\pi ft) dt,$$

R_{NN} being the auto-correlation function of the regions of the original extracted ion chromatograms having no chromatographic
30 peaks.

11. The method of claim 7 wherein the noise reduction step includes performing matched filtration of the original extracted

ion chromatograms for a full range of mass-to-charge (m/z) values in the mass spectral dimension of the first data array, and determining at least one transfer function for matched filtration for a predetermined number of the m/z values.

5

12. The method of claim 4 wherein the second analyzing step includes detecting peak candidates in the chromatographic time dimension for a full range of mass-to-charge (m/z) values in the mass spectral dimension, and calculating a respective peak score value for each detected peak candidate.

10

13. The method of claim 12 wherein the second analyzing step includes analyzing the second data array to determine the presence of one or more chromatographic peaks within the second data set based on a comparison of the calculated peak score values with the detected peak candidates.

15

14. The method of claim 12 wherein the calculating step includes calculating the respective peak score value for each detected peak candidate based on a plurality of criteria selected from the group consisting of the shape, the width, and the intensity of the detected peak candidate.

20

15. The method of claim 14 wherein the calculating step includes calculating the respective peak score value for each detected peak candidate based on at least one ratio of peak candidate intensities.

25

16. The method of claim 15 wherein the calculating step includes calculating each respective peak score value Sc_f according to the equation

30

$$Sc_f = Sc \cdot K_v \cdot K_I,$$

Sc being a first peak score value determined by examining the detected peak candidate in the chromatographic time dimension, and K_v and K_t being second and third peak score values, respectively, the second and third peak score values being determined by examining the shape of the detected peak candidate in the mass spectral dimension and by examining at least one ratio of intensities of peak candidates in at least one isotopic cluster in the mass spectral dimension.

10

17. The method of claim 16 further including the step of calculating the first peak score value Sc as a ratio of a maximum intensity of the noise-reduced extracted ion chromatograms to a mean intensity of the noise-reduced extracted ion chromatograms, a value of Sc greater than a first predetermined value being indicative of the presence of a peak, and a value of Sc less than a second predetermined value being indicative of the absence of peaks.

15

18. The method of claim 16 wherein the calculating step includes comparing each respective peak score value Sc_f to a predetermined peak score threshold value Sc_t to detect a first peak in each chromatogram.

20

19. The method of claim 18 further including the step of eliminating the first detected peak from the respective extracted ion chromatogram, and performing the noise reducing step, the second generating step, and the second analyzing step to detect a second peak in each chromatogram.

25

30

20. The method of claim 19 further including the step of repeating the eliminating step, the noise reducing step, the second generating step, and the second analyzing step a

predetermined number of times equal to an expected number of peaks in each chromatogram.

21. The method of claim 4 further including comparing chromatographic elution profiles of two or more compounds of the mixture.

22. The method of claim 21 further including calculating a cross-correlation function of the noise-reduced extracted ion chromatograms corresponding to two or more co-eluting compounds of the mixture, the cross-correlation function having associated maximum and mean values, wherein a predetermined high value of a ratio of the maximum value to the mean value is indicative of similar chromatographic elution profiles of the co-eluting compounds, thereby enabling natural isotopes or isotopically-labeled components having the same elution profile as a mono-isotopic peak to be distinguished from different components having similar elution times and different chromatographic elution profiles.

23. A method of analyzing a multi-dimensional data set to determine the presence of one or more peaks within the data set, the peaks being representative of respective compounds, used for detecting at least one constituent compound in a sample mixture, comprising the steps of:

generating a first multi-dimensional data set, the first data set being representable by a first data array including a first dimension corresponding to compound separation and a second dimension corresponding to compound characteristics;

analyzing the first data array in the compound separation dimension to generate one or more original representations of the compound separation, the original representations including one or more peaks;

detecting one or more peak candidates in the original representations of the compound separation; and

analyzing the first data array in the compound characteristics dimension to determine the presence of one or more peaks within the data set based on a comparison of at least one characteristic of the respective peak candidates.

24. The method of claim 23 wherein the peaks representative of respective compounds are selected from the group consisting of chromatographic peaks, electro-chromatographic peaks, and electrophoretic peaks.

25. The method of claim 23 further including determining noise characteristics of the first data set based on an analysis of one or more regions of the original representations having no peaks, and reducing noise in the first data set by performing a noise reduction technique on the original representations of the compound separation based on the determined noise characteristics, thereby generating one or more noise-reduced representations of the compound separation, the noise-reduced representations including one or more peaks.

26. The method of claim 25 further including generating a second multi-dimensional data set with reduced noise based on the noise-reduced representations, the second data set being representable by a second data array including a compound separation dimension and a compound characteristics dimension, and analyzing the second data array in the compound separation dimension and the compound characteristics dimension to determine the presence of at least one peak within the second data set.

27. The method of claim 25 wherein the original representations of the compound separation generated in the first analyzing step

and the noise-reduced representations of the compound separation generated in the noise reduction step are selected from the group consisting of an extracted ion chromatogram, an electrochromatogram, and an electropherogram.

5

28. A method of quantitating a level of chromatographic peak overlapping in a tandem MS (MS/MS) mode based on the analysis of LC-MS data, used to increase the number of positive identifications of constituent compounds in the sample mixture, comprising the steps of:

10

generating a multi-dimensional data set, the data set being representable by a data array including a first dimension corresponding to compound separation and a second dimension corresponding to compound characteristics;

15

analyzing the data array in the compound separation dimension and the compound characteristics dimension to generate one or more representations of the compound separation, the representations including a plurality of chromatographic peaks and a plurality of MS/MS wells, each chromatographic peak being representative of a respective compound, wherein one or more of the MS/MS wells have two or more overlapping chromatographic peaks associated therewith;

20

identifying a plurality of overlapping chromatographic peaks in the representations of the compound separation;

25

quantitating levels of overlapping for the respective overlapping chromatographic peaks; and

selecting one or more of the MS/MS wells with not more than two overlapping chromatographic peaks of commensurate intensity in one well to increase the number of chromatographic peaks for positive identification.

30

29. The method of claim 28 wherein the quantitating step includes quantitating levels of overlapping for the respective overlapping chromatographic peaks according to

5
$$\text{ovlp} = \sum_{i=1}^M C_i \cdot \exp(-\Delta t_i^2 / 2\sigma_i^2) / C_0,$$

wherein M is the number of chromatographic peaks within an MS/MS window having an intensity greater a predetermined portion of a
10 main chromatographic peak, C_i is the intensity of the overlapping chromatographic peaks, Δt_i^2 is a distance between the main chromatographic peak and the overlapping chromatographic peak in the compound separation dimension, and σ_i is the compound separation variance.

15 30. A system for analyzing a multi-dimensional data set to determine the presence of one or more peaks within the data set, the peaks being representative of respective compounds, used for detecting at least one constituent compound in a sample mixture,
20 comprising:

a compound-separating unit configured to separate the constituent compounds in the sample mixture, and to generate data corresponding to the separated compounds;

25 a compound-analyzing unit configured to analyze the separated compounds and to generate data corresponding to characteristics thereof; and

a computer operative to:

30 generate a first multi-dimensional data set, the first data set being representable by a first data array including a first dimension corresponding to compound separation and a second dimension corresponding to compound characteristics;

analyze the first data array in the compound separation dimension to generate one or more original representations of the

compound separation, the original representations including one or more peaks and one or more regions having no peaks;

determine noise characteristics of the first data set based on an analysis of the regions of the original representations

5 having no peaks;

reduce noise in the multi-dimensional data set by performing matched filtration of the original representations of the compound separation with at least one transfer function based on the determined noise characteristics, thereby generating one or more
10 noise-reduced representations of the compound separation, the noise-reduced representations including one or more peaks;

generate a second multi-dimensional data set with reduced noise based on the noise-reduced representations, the second data set being representable by a second data array including a
15 compound separation dimension and a compound characteristics dimension; and

analyze the second data array in the compound separation dimension and the compound characteristics dimension to determine the presence of one or more peaks within the second data set.

20

31. The method of claim 30 wherein the original and noise-reduced representations of the compound separation generated by the computer are selected from the group consisting of an extracted ion chromatogram, an electro-chromatogram, and an
25 electropherogram.

32. The method of claim 30 wherein the peaks representative of respective compounds are selected from the group consisting of chromatographic peaks, electro-chromatographic peaks, and
30 electrophoretic peaks.